TITLE:
NP完全集合によるcoNP集合の近似とその応用について

AUTHOR(S):
宮崎, 修一; 岩間, 一雄

CITATION:
宮崎, 修一 ...[et al]. NP完全集合によるcoNP集合の近似とその応用について. 電子情報通信学会論文誌 1998, J81-D1(6): 677-684

ISSUE DATE:
1998-06-25

URL:
http://hdl.handle.net/2433/227031

RIGHT:
© 1998 電子情報通信学会(IEICE)

京都大学学術情報リポジトリ
KURENAI 紅
Kyoto University Research Information Repository

京都大学 KYOTO UNIVERSITY
A Self-archived copy in
Kyoto University Research Information Repository
https://repository.kulib.kyoto-u.ac.jp
京都大学学術情報リポジトリ KURENAI 紅 Kyoto University Research Information Repository

NP完全集合によるcoNP集合の近似とその応用について

宮崎 修一† 岩間 一雄†

Approximation of coNP Sets by NP-Complete Sets and Its Application

Shuichi MIYAZAKI† and Kazuo IWAMA†

あらまし クラス C_1 に属する集合 L_1 が，クラス C_2 に属する集合 L_2 の部分集合であるとき，L_1 は L_2 の近似であると言う．$L_1 \subset L'_1$ であり，$L'_1 - L_1$ が無限集合となるような近似 L'_1 が存在しないとき，L_1 は最適近似であると言う．C_1 がクラス P，C_2 がクラス NP である場合には，P$\neq$ NP ならば弱い条件のもとで最適な近似が存在しないことが示されている．本論文では，C_1 が NP 完全集合のクラスで，C_2 が coNP である場合に対し，同様の結果を示す．coNP 集合の NP 完全集合による近似は，組合せアルゴリズムを実験的に評価する際の例題生成の効率に深いかかわりをもっている．

キーワード 近似，NP 完全集合，coNP 集合，例題生成

1. まえがき

集合 L_1，L_2 がそれぞれクラス C_1，C_2 に属するとする．集合 L_1 が集合 L_2 の部分集合であるとき，L_1 は L_2 の近似である（または，L_1 は L_2 を近似する）と言う．二つの集合 A，B が，ともに集合 L の近似で，$B \subset A$ であり，$A - B$ が無限集合であるとき，集合 A は集合 B よりも良い近似であると言う．集合 A よりも良い近似が存在しないとき，集合 A は最適な近似であると言う．複雑な組合せ問題に対する良いアルゴリズムを構築することは，（例えばクラス P のような）やさしいクラスに属する良い近似を求めることと同じである．例えば CNF 論理式の充足可能性問題（SAT）を考えたとき，より多くの例題に対して多項式時間で動作するアルゴリズムを開発することは，クラス P に属する，より大きな *SAT*（充足可能な CNF 論理式の集合）の部分集合を見つけることと同じである．

また，組合せ判定アルゴリズムの性能評価のための例題生成においても，近似の概念が重要な意味をもつ．近年では，（yes，no の）答を入力として与え，確率アルゴリズムを用いて例題を生成するという方法が主流である．例えば SAT の場合，例題生成アルゴリズムは，*SAT*および，*UNSAT*（充足不能な CNF 論理式の集合）をそれぞれ生成する二つの独立なアルゴリズムにより構成されている．例題生成アルゴリズムは基本的には非決定性のチューリング機械であるため，*SAT*のような NP 言語を多項式時間で生成することは可能である．ところが，NP$\neq$coNP ならば，*UNSAT*等の coNP 言語を多項式時間で生成することは不可能である．従って，多項式時間で動作する例題生成アルゴリズムを開発するために，coNP 集合の NP 集合による近似を見つけることが必要となってくる．

上述の場合，できるだけ大きな近似を見つけることは，できるだけ多くの例題を多項式時間で生成できることを意味する．従って，この場合も良い近似を見つけることが重要となってくる．ところが，例題生成の場合は，クラス P のようなやさしいクラスの近似を見つけるのでは意味をなさない．例えば，次のような，充足不能な論理式を生成するアルゴリズムを考えてみる．(i) 任意の論理式 f を生成する．(ii) 任意に二つの変数 x_i，x_j を選び，f に四つの項 $(x_i + x_j)(x_i + \overline{x_j})(\overline{x_i} + x_j)(\overline{x_i} + \overline{x_j})$ を付け加える．このアルゴリズムは，確かに *UNSAT*の部分集合を生成する．しかし，生成される論理式からなる集合はクラス P に属するため，SAT アルゴリズムの性能評価には役立たない．

近年，充足不能な CNF 論理式 [1], [8]，非ハミルトングラフ [9]，k 色塗分け不能なグラフ [11] といった

† 京都大学大学院情報学研究科，京都市
Graduate School of Informatics, Kyoto University, Yoshida-Honmachi, Sakyo-ku, Kyoto-shi, 606–8501 Japan

京都大学 KYOTO UNIVERSITY
A Self-archived copy in
Kyoto University Research Information Repository
https://repository.kulib.kyoto-u.ac.jp
京都大学学術情報リポジトリ KURENAI 紅 Kyoto University Research Information Repository

coNP 言語の例題生成アルゴリズムが提案されているが，それらは上述の近似の概念を利用したものである．中でも，AIM ジェネレータ [1] は，実際にベンチマーク例題を生成するのに稼動している．AIM ジェネレータは次のような特徴をもっている [8]．無限の集合列 $L_0, L_1, \cdots, L_i, L_{i+1}, \cdots$ が存在し，(i) 各 L_i は *UNSAT* の近似である．(ii) 各 L_i は NP 完全である．(iii) $L_i \subset L_{i+1}$ であり，$L_{i+1} - L_i$ は $\mathrm{D^P}$ 完全である（すなわち L_{i+1} は本質的に L_i よりも良い近似である）．直観的には AIM ジェネレータは *UNSAT* のいくらでも良い近似を生成できるということになる．

それでは，いくらでも大きな近似が存在するのか，それとも最適な近似が存在するのかという疑問が当然生じる．NP 集合を P 集合で近似する場合には，NP 集合 A がある弱い条件を満たせば，A には最適な P 近似は存在しないという結果（P-levelability）が示されている [10]．この文献では，NP 集合 A が従順充填集合ならば，(i) A にはクラス P に属する近似 B が存在し，B は無限集合である，(ii) P$\neq$ NP ならば，A は最適な P 近似をもたない，ということが示されている [10]．

本論文では，coNP 集合を NP 完全集合で近似する場合に対し，同様の議論を行う．本論文では，(i) coNP 完全集合 A は，ある弱い条件の元で NP 完全近似をもつ，(ii) NP$\neq$coNP であり，coNP 集合 A，および A の NP 完全近似 B がある条件を満たすならば，B よりも良い NP 完全近似が存在する，ということを示す．[10] では，NP 集合 A の P 近似 B が最適でないためには，A に対する条件だけが必要であったが，我々の結果 (ii) においては，NP 完全近似 B が最適でないためには，A，B 両方の集合に条件が必要である．すなわち，我々の条件は [10] の条件よりも強いものである．しかし，我々の条件も十分に一般的である．**4.** で，条件の一般性を示すために，coNP 完全集合 *NH* と現在知られている最も良い NP 完全近似である *NS2F* が我々の条件を満たすことを示す．更に，**5.** では，我々の条件を弱めることができることを示す．

2. 定義および背景

本論文では，集合は $\{0,1\}^*$ の部分集合であり，関数は $\{0,1\}^*$ から $\{0,1\}^*$ への写像を意味する．列 f の長さを $|f|$ で，集合 S の濃度を $|S|$ で表す．集合 $\{\psi(f) \mid f \in S\}$ のことを単に $\psi(S)$ と書く．任意の列 $f \in S$ に対して，$\psi(f) \in S$ が成り立つとき，集合 S は関数 ψ で閉じていると言う．任意の列 f に対して，$|\psi(f)| > |f|$ であるとき，関数 ψ は長さ増加関数であると言う．任意の列 f，g に対して，$\delta(f) = \delta(g)$ ならば $f = g$ であるとき，δ は 1 対 1 関数であると言う．任意の集合 S_1，S_2 に対し，$\delta(S_1) \cap \psi(S_2) = \emptyset$ であるとき，関数 δ と ψ は互いに素であると言う．ある集合の P（NP，NP 完全）集合による近似を単に P（NP，NP 完全）近似と書く．

まずはじめに，NP 集合を P 集合で近似する場合について述べる [2]．

［命題 1］ A は $A \notin \mathrm{P}$ である集合とする．また，次の条件を満たす関数 ψ が存在するとする．

（i） 集合 A および $\overline{A}$ は ψ で閉じている．
（ii） ψ は長さ増加関数である．
（iii） ψ は多項式時間で計算可能である．

このとき，集合 A には最適な P 近似は存在しない．

（証明） [2] 最適な P 近似 B が存在したと仮定する．このとき，集合 C を以下のように定義する．

$$C = \{f \mid f \notin B \text{ かつ } \psi(f) \in B\}.$$

$B \in \mathrm{P}$ で，ψ が多項式時間で計算可能なので，$C \in \mathrm{P}$ である．各 $f \in A$ に対して，以下のような無限集合 D_f を考える．

$$D_f = \{f, \psi(f), \psi^2(f), \cdots, \psi^n(f), \cdots\}.$$

集合 A は ψ で閉じているので，$D_f \subseteq A$ である．ψ は長さ増加関数であるので，集合 D_f はクラス P に属する．また，B は最適近似であるので，$D_f - B$ は有限集合である．もし，$D_f - B$ が無限集合なら，$B \cup D_f$ が B よりも良い P 近似であり，仮定に矛盾するからである．

$A \notin \mathrm{P}$ なので，$A - B$ は無限集合である．$B \cap D_f \neq \emptyset$ なので，各 $f \in A - B$ に対して，集合 D_f は集合 C のある列 g を含んでいる．また，ψ は長さ増加関数なので $|g| \geq |f|$ である．($|g| = |f|$ であるのは，$g = f$ のときのみである．) 従って，ある固定された g に対して，$g \in D_f$ となる $f \in A - B$ の数は有限である．$A - B$ は無限集合であるので，上述のような列 g は無限に存在することになる．よって，集合 C は無限集合である．集合 C の定義より，$C \cap B = \emptyset$ であるので，集合 $C - B$ は無限集合である．また，$B \in \mathrm{P}$，$C \in \mathrm{P}$ なので，$B \cup C \in \mathrm{P}$ である．従って，集合 $B \cup C$ は集合 B よりも良い近似であり，集合 B が最適近似であるという仮定に矛盾する． □

京都大学 KYOTO UNIVERSITY
A Self-archived copy in
Kyoto University Research Information Repository
https://repository.kulib.kyoto-u.ac.jp
京都大学学術情報リポジトリ KURENAI 紅 Kyoto University Research Information Repository

ψ が長さ増加関数であるという条件を弱めることができる．関数 ψ が次の条件を満たすとき，ψ は逆関数が有限停止であると言う．(i) すべての列 f に対して $|\psi^{-1}(f)|$ が有限である．但し，$\psi^{-k}(f)=\{g \mid f=\psi^k(g)\}$ である．(ii) ある多項式 $p(x)$ が存在して，すべての列 f に対して $\psi^{-p(|f|)}(f)=\phi$ である．命題1において，ψ が長さ増加関数である代わりに ψ は逆関数が有限停止であるとしても次の二つの条件は満たされる．

（a） 集合 D_f は無限集合でクラス P に属する．

（b） ある g に対し，$g \in D_f$ となる f の個数は有限なので，集合 C は無限集合である．

従って，命題1は逆関数が有限停止である ψ に対しても成り立つことがわかる．以下，命題2の ψ，定理1の ψ，δ，および定理3の ψ，δ はすべて，長さ増加の条件を逆関数が有限停止であるに変更しても成立する．

命題1と同様にして NP 集合による近似の場合も容易に証明できる．

［命題 2］ A は $A \notin \mathrm{NP}$ である集合とする．また，多項式時間計算可能な長さ増加関数 ψ が存在し，A および $\overline{A}$ は ψ で閉じているとする．このとき，集合 A には最適な NP 近似は存在しない．

（証明） 最適な NP 近似 B が存在したとする．集合 C を $C=\{f \mid \psi(f) \in B\}$ と定義すると，$C \in \mathrm{NP}$ であり，$C \cup B \in \mathrm{NP}$ である．この場合，C と B が共通部分をもってしまう可能性があるが，この場合も命題1と同様に，$C-B$ は無限集合であり，$C \cup B$ が B よりも良い近似であるということが言える． □

P-levelability の定義には，最適な P 近似が存在しないことだけでなく，無限集合である P 近似が存在することも含まれている．命題1の場合には，ψ が長さ増加関数なので，無限集合近似の存在は以下のようにして容易に示すことができる．任意の列 $f \in A$ に対して，集合 $\{f, \psi(f), \psi^2(f), \cdots\}$ は無限集合であり，クラス P に属する．ところが，[10] では，近似の存在が自明でない場合についても述べられている．

最適近似の存在を考える前に，まず，近似の存在，すなわち coNP 完全集合に NP 完全近似が存在するかどうかについて考えてみる．例えば Berman-Hartmanis 予想 [4] が成り立つならば，この主張は明らかに成り立つ．しかし，実際は，これよりもより弱い条件のもとでも我々の主張が成り立つ．

［命題 3］ L を coNP 完全集合とする．*UNSAT*から L への多項式時間変換 α が存在し，α は1対1，かつ，長さ増加関数であるとする．このとき，L は NP 完全近似をもつ．

（証明） 明らかに *UNSAT*は NP 完全近似 R をもつ（例えば，$n-1$ 変数からなる充足可能な論理式に $(x_n)(\overline{x_n})$ を加えてできる論理式の集合や，AIM ジェネレータ [1] によって生成される集合等を R と考えることができる)．α は *UNSAT*から L への多項式時間変換なので，$\alpha(UNSAT) \subseteqq L$ である．また，$R \subseteqq UNSAT$ なので，$\alpha(R) \subseteqq \alpha(UNSAT)$ である．従って，$\alpha(R) \subseteqq L$ となる．

R が NP 完全集合であり，α は1対1関数なので，$\alpha(R)$ は NP 困難である．また，$\alpha(R)$ が NP 集合であることは，以下のような非決定性チューリング機械 M の動作を考えれば容易にわかる．列 y が与えられたとき，M は列 $x \in R$ を非決定的にゲスし，$\alpha(x)$ を計算し，$\alpha(x)=y$ かどうかを検証する．α が長さ増加関数なので，この動作は多項式時間で行うことができる（α が長さ増加関数であるという条件は，ある多項式 p が存在し，すべての x に対して $|x| \leqq p(|\alpha(x)|)$ であるという条件に弱めることができる)．従って，$\alpha(R)$ は L の近似であり，NP 完全集合である． □

最後に，近似の概念を利用した例題生成法について，SAT を例にとり簡単に述べる [1]．以下のアルゴリズム [1] は，N 項からなる充足可能なすべての論理式を（非決定性の）多項式時間で生成することができる．

（1） 解 $(a_1, a_2, \cdots, a_n)$ を非決定的に選ぶ．但し，a_i は0または1である．

（2） （1）で選んだ割当てのもとで0とならないような項 C を非決定的に生成する．このことは，変数 x_i をランダムに選び，$a_i=1$ ならば x_i を，$a_i=0$ なら $\overline{x_i}$ をあらかじめ項 C に加えることにより行うことができる．

（3） （2）を N 回繰り返す．

このように，NP 集合を生成するのは比較的簡単である．しかし，coNP 集合の生成は効率的に行うことができない．充足不能な論理式を生成する以下のアルゴリズム [1] を考えてみる．

（1） 変数 x_i をランダムに選び，$f=(x_i)(\overline{x_i})$ とする．

（2） 以下の四つの規則の中から，適用できるものをランダムに選び f に適用させる．但し，x はリテラルを，A はリテラルの和を表す．

（i） $(A) \rightarrow (A+x)(A+\overline{x})$,

（ii） $(A)(A+x) \rightarrow (A)$,

京都大学 KYOTO UNIVERSITY
A Self-archived copy in
Kyoto University Research Information Repository
https://repository.kulib.kyoto-u.ac.jp
京都大学学術情報リポジトリ KURENAI 紅
Kyoto University Research Information Repository

(iii) $(A+x) \to (A)$,

(iv) ランダムな項を f に加える.

(3) f を出力する.または,(2) に戻る.

このアルゴリズムは,Resolution と呼ばれる定理証明系の規則を逆に適用したものである.Resolution は,すべての充足不能式を証明できるので,このアルゴリズムはすべての充足不能式を生成することができる.しかし,Resolution での証明が指数ステップかかる論理式の集合が存在するため[6],このアルゴリズムは多項式時間では動作しない.このアルゴリズムが多項式時間で動作するように改良するために,(ii) の規則の適用回数を例えば定数回に制限することにより,*UNSAT*の部分集合を多項式時間で出力させるようにすることができる.すなわち,近似の概念を利用するわけである.(ii) の規則を全く使わなくても,生成される例題集合は NP 完全であるということが証明されている [1],[8].

3. 主 定 理

前章で見たように,最適な P 近似が存在しないことの証明を,最適な NP 近似が存在しないことの証明に応用するのは簡単である.そこで,最適な NP 完全近似の存在について考えてみる.集合 B が集合 A の NP 完全近似であるとする (図 1).B は NP 完全集合なので,任意の NP 完全集合 X に対して,ある多項式時間変換 α が存在し,$\alpha(X) \subseteq B$,$\alpha(\overline{X}) \subseteq \overline{B}$ となる.よって,$A-B$ が無限集合であっても,その大部分が $\alpha(\overline{X})$ と重なる可能性がある.この場合,命題 2 で用いた,集合 B を拡張するという方法は使うことができない (新しく拡張された集合が $\alpha(\overline{X})$ と重なってしまうと,その集合の NP 完全性が保証できないからである).このように,命題 2 を NP 完全近似に応用するのは容易ではない.

主定理には,二つの関数 ψ および δ を用いる.ψ,δ は命題 1,2 と同様に,多項式時間計算可能な長さ増加関数である.更に,δ は以下の三つの条件を満たす関数である (4. に具体的な ψ,δ の例を述べる).

(i) δ は 1 対 1 である.

(ii) δ^{-1} は多項式時間で計算可能である.

(iii) δ と ψ は互いに素である.

[定理 1] A を coNP 集合,B を A の NP 完全近似であるとする.ψ,δ は上述の条件を満たす関数であるとする.A,$\overline{A}$ は,ψ,δ で閉じているとする.また,$\overline{B}$ は δ で閉じているとする.このとき,NP $\neq$ coNP ならば,B は最適近似ではない.

(証明) B は NP 完全集合なので,任意の NP 完全集合 X に対して,ある多項式時間変換 α が存在し,$\alpha(X) \subseteq B$,$\alpha(\overline{X}) \subseteq \overline{B}$ である.まず,$\alpha(\overline{X}) \subseteq \overline{A}$ となるような X および α が存在する場合について考える.この場合は,命題 2 の方法をそのまま適用しても,集合 B を拡張したとき $\alpha(\overline{X})$ と重なることがないため,命題 2 と同様に証明できる.従って,今後は,どのような (X,α) の組も $\alpha(\overline{X}) \cap A \neq \emptyset$ を満たすと仮定する.

まず,以下のような表記を定義する.$g=\delta(f)$ となる列 f が存在しないとき,$\delta^{-1}(g)=\Lambda$ と書く.また,$\delta^{-1}(g)=\Lambda$ であり,ある $i \geqq 0$ が存在して $\delta^{i}(g)=f$ となるとき $g=\delta^{-\infty}(f)$ と書く.

集合 C を $C=\alpha(\overline{X}) \cap A$,$C_0$ を $C_0=\{\delta^{-\infty}(f) \mid f \in C\}$ と定義する (図 2).C_0 は以下の理由により,無限集合であることがわかる.C_0 が有限集合であると仮定する.すると,列 $f \in \alpha(X \cup \overline{X})$ が与えられたとき,$f \in C$ かどうかを多項式時間で判定することができる.これは,f に δ^{-1} を適用できなくなるまで繰り返し適用させ,最後に得られた列が C_0 に含まれるかどうかをチェックすることにより行うことができる.

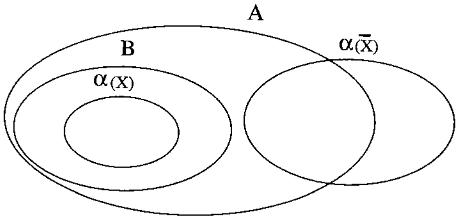


図 1 NP 完全近似
Fig. 1 An NPC-approximation.

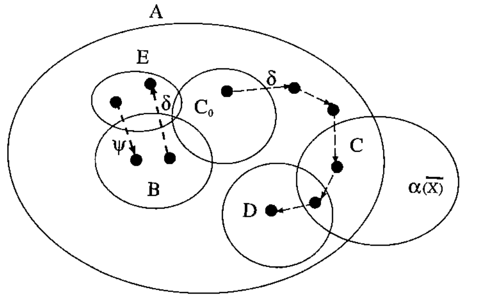


図 2 定理 1 の証明
Fig. 2 Proof of Theorem 1.

京都大学 KYOTO UNIVERSITY
A Self-archived copy in
Kyoto University Research Information Repository
https://repository.kulib.kyoto-u.ac.jp
京都大学学術情報リポジトリ KURENAI 紅 Kyoto University Research Information Repository

そこで，多項式時間変換 α から新たな多項式時間変換 α' を以下のように定義する．列 x が与えられたとき，$\alpha(x) \notin C$ なら $\alpha'(x) = \alpha(x)$，$\alpha(x) \in C$ なら $\alpha'(x)$ は $\overline{A}$ のある列とする．上述のように，$\alpha(x)$ が C に含まれるかどうかを多項式時間で判定できるので，α' は多項式時間変換である．ところが $\alpha'(\overline{X}) \cap A = \emptyset$ であり，そのような X，α' の組が存在しないという仮定に反する．従って C_0 は無限集合である．

次に $D = \delta(C)$，$E = \{f \mid \text{(a) } \delta^{-1}(f) \in B$ または (b) $\delta^{-1}(f) = \Lambda, \psi(f) \in B\}$ と定義する．

ここで，以下の（i）～（v）について考える．

（i） 集合 A は δ，ψ^{-1} で閉じているので，$E \subseteq A$ である．

（ii） $B \in \mathrm{NP}$ で，δ^{-1} と ψ は多項式時間計算可能なので，$E \in \mathrm{NP}$ である．

（iii） 以下の理由により $E \cap D = \emptyset$ である．$f \in D$ とすると，定義より $\delta^{-1}(f) \in C$ である．よって集合 E の定義の条件 (b) は f には当てはまらない．また，$B \cap C = \emptyset$ なので，$\delta^{-1}(f) \notin B$ である．従って条件 (a) も当てはまらず，$E \cap D = \emptyset$ が言える．

（iv） 以下の理由により $C - D$ は無限集合である．ある列 $f_0 \in C_0$ に δ を繰り返し適用することを考える．$f_{i+1} = \delta(f_i)$ $(i \geqq 0)$ とする（図 2 参照）．$f_0, f_1, f_2, \cdots$ の列の中で，最初に C に入る要素を f_i とする．集合 D の定義より，$f_i, f_{i+1}, \cdots, f_j$ が C に含まれているなら，$f_{i+1}, \cdots, f_{j+1}$ は D に含まれている．つまり，f_i は $C - D$ に含まれている．同様に f_0 と異なる列 $g_0 \in C_0$ に対しても同様に δ を適用させる．δ は 1 対 1 関数なので，$f_i = g_j$ となるような i, j は存在しない．すなわち，C_0 中の異なる二つの列 f_0，g_0 に δ を適用させて，$C - D$ の同じ要素に到達することはあり得ない．C_0 は無限集合なので $C - D$ も無限集合である．

（v） 以下の理由により，$E - B$ は無限集合である．各 $f \in C - D$ に δ^{-1} を繰り返し適用させていくと，C_0 の定義より，最後には C_0 の要素が得られる．この過程には次のような二つの場合が考えられる．(a) この過程の途中で，いずれかの要素が集合 B に含まれる場合，および，(b) どの要素も B に含まれることなく C_0 の要素に到達する場合．(a) の場合，初めて B に含まれる直前の要素を g_1 とする．すなわち，$\delta^{-1}(g_1) \in B$ であり，f から g_1 に至るまでにはどの列も B には含まれていない．また，(b) の場合，最後に得られた C_0 の要素を g_2 とする．すなわち，$\delta^{-1}(g_2) = \Lambda$ である．（iv）より $C - D$ は無限集合で，δ が 1 対 1 関数なので，上述のような g_1 および g_2 から構成される集合も無限集合である．定義から $g_1 \in E - B$ なので，g_1 が無限に存在すれば，$E - B$ は無限集合である．また，g_2 が無限に存在するなら，命題 1，2 と同様の議論で，g_2 に ψ を繰り返し適用していくと，$E - B$ が無限集合であるということが言える．この場合，δ と ψ は互いに素なので，$\{g_2, \psi(g_2), \psi^2(g_2), \cdots\}$ という集合は D と重ならないことに注意されたい．

（i）より，集合 $E \cup B$ は集合 A の近似であり，(ii) より，$E \cup B \in \mathrm{NP}$ である．また，(v) より，$(E \cup B) - B$ は無限集合である．最後に，$E \cup B$ が NP 完全集合であることを示す．そのために，多項式時間変換 α' を $\alpha' = \delta \cdot \alpha$ と定義する．α' は明らかに多項式時間計算可能である．$\alpha'(\overline{X})(= \delta(\alpha(\overline{X})))$ は D および，$\overline{A}$ の要素のみを含むため，上述の（iii）より，E とは共通部分をもたない．また，$\overline{B}$ が δ で閉じているので，$\alpha'(\overline{X})$ は B とも共通部分をもたない．よって $\alpha'(\overline{X}) \subseteq \overline{E \cup B}$ である．また，すべての $g \in B$ に対して，E は $\delta(g)$ を含んでいるので，$\alpha'(X) \subseteq E \cup B$ である．すなわち，$E \cup B$ は NP 完全集合であり，B よりも良い近似である．□

我々は例題生成への応用を考えているため，定理 1 において，集合 A は coNP 集合としているが，実際には NP 以外のクラスの集合であれば定理は成り立つ．また，定理 1 における，$\overline{B}$ が δ で閉じているという条件は弱めることができる（5. 参照）．

4. 主定理の応用

本章では，具体的な coNP 完全集合とその NP 完全近似に対して，定理 1 の条件を満たす関数 δ と ψ を示す．coNP 完全集合としては，ハミルトン閉路をもたないグラフからなる集合 *NH* を，また，*NH* の NP 完全近似として，非 *sub-2-factor* であるグラフの集合 *NS2F* [5] を取り扱う．*NH* には，*N1T* [3] という NP 完全近似が知られているが，*NS2F* は *N1T* よりも良い NP 完全近似であることが示された [9]．

NS2F は，頂点集合 V を，（i）$T \neq V$，（ii）$w(T) > |S| + \sum_{Q_i} \left\lfloor \frac{edge(Q_i, T)}{2} \right\rfloor$ を満たすような互いに素な R，S，T $(V = R \cup S \cup T)$ に分割することのできるグラフの集合である．ここで，$w(T)$ は T の連結成分の数，Q_i は R の各連結成分，$edge(Q_i, T)$ は Q_i の頂点と T の頂点を結ぶ枝の数である．図 3 に *NS2F* のグラフの例を示す．図に示された頂点分割では，$w(T) = 4$，

京都大学 KYOTO UNIVERSITY
A Self-archived copy in
Kyoto University Research Information Repository
https://repository.kulib.kyoto-u.ac.jp
京都大学学術情報リポジトリ KURENAI 紅 Kyoto University Research Information Repository

$|S|=1$，$\sum_{Q_i}\left\lfloor\frac{edge(Q_i,T)}{2}\right\rfloor=\left\lfloor\frac{3}{2}\right\rfloor+\left\lfloor\frac{3}{2}\right\rfloor=2$ となっており，上述の条件（i），(ii) を満たしていることがわかる．

ここで，定理 1 の条件について再び触れておく．δ と ψ の満たすべき条件は以下のとおりである．(i) δ と ψ は多項式時間計算可能な長さ増加関数である．(ii) δ は 1 対 1 である．(iii) δ^{-1} は多項式時間で計算可能である．(iv) δ と ψ は互いに素である．(v) A と $\overline{A}$ は δ と ψ で閉じており，$\overline{B}$ は δ で閉じている．ここで，A は coNP 集合を，B は A の NP 完全近似を表す．

関数 δ と ψ を図 4 に示す．形式的には，グラフ $G=(V,E)$ $(V=\{v_1,v_2\cdots,v_{n-1},v_n\})$ に対して，$\delta(G)=(V',E')$，$(V'=V\cup\{v_{n+1},v_{n+2}\}$，$E'=E\cup\{(v_n,v_{n+1}),(v_{n+1},v_{n+2})\}\cup\{(v_{n+2},v_i)\mid(v_n,v_i)\in E\})$，$\psi(G)=(V'',E'')$ $(V''=V\cup\{v_{n+1},v_{n+2},v_{n+3}\},E''=E\cup\{(v_n,v_{n+1}),(v_{n+1},v_{n+2}),(v_{n+2},v_{n+3})\}\cup\{(v_{n+3},v_i)\mid(v_n,v_i)\in E\})$ である．

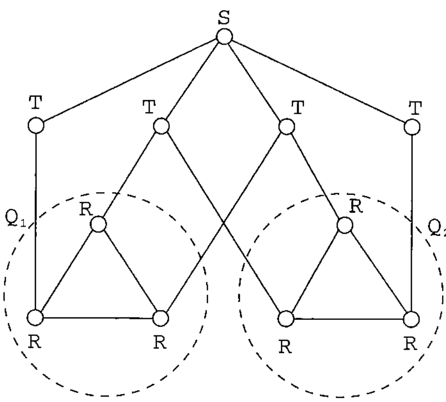


図 3　NS2F のグラフの例
Fig. 3　A graph in NS2F.

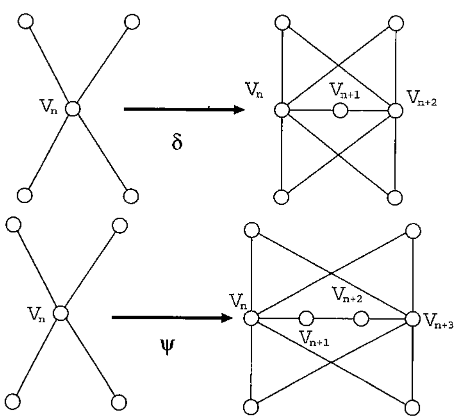


図 4　関数 δ および ψ
Fig. 4　Functions δ and ψ.

［定理 2］　δ と ψ は条件（i）～（v）を満たしている．

（証明）　本証明では $\overline{B}$ が δ で閉じていることのみを示す．他の条件が満たされていることは明らかである．このことを示すためには，任意のグラフ G に対して，$\delta(G)\in$NS2F ならば $G\in$NS2F となることを示せばよい．$\delta(G)\in$NS2F であると仮定する．すると，$\delta(G)$ の頂点集合 $V'(=V\cup\{v_{n+1},v_{n+2}\})$ を R'，S'，T' に分割する Π' が存在し，

$$w(T')>|S'|+\sum_{Q'_i}\left\lfloor\frac{edge(Q'_i,T')}{2}\right\rfloor \tag{1}$$

を満たしている．以下，

$$w(T)>|S|+\sum_{Q_i}\left\lfloor\frac{edge(Q_i,T)}{2}\right\rfloor \tag{2}$$

を満たす，G の頂点集合 V から R，S，T への分割 Π が存在することを示す．v_n，v_{n+1}，v_{n+2} が Π' によってどのように分割されているかに応じて，以下の三つの場合に分けて考える．

(A)　v_n，v_{n+1}，v_{n+2} のうち，少なくとも一つが S' に入っている場合．

(B)　v_n，v_{n+1}，v_{n+2} のうち少なくとも二つが R' に入っていて，残りは T' に入っている場合．

(C)　v_n，v_{n+1}，v_{n+2} のうち R' に入っているものはたかだか一つで，残りが T' に入っている場合．

場合（A）：まず，Π' の元で $v_{n+1}\in S'$ または $v_{n+2}\in S'$ となっている場合を考える．この場合は，Π を Π' と全く同じようにする．すなわち，グラフ $\delta(G)$ の頂点 v_i が Π' の元で R' に入っているなら，Π はグラフ G の頂点 v_i を R に割り当てるという具合である．v_{n+1}，v_{n+2} のうち，T' に入っているものはたかだか一つなので，v_{n+1} と v_{n+2} を取り去ることによって $w(T')$ はたかだか 1 しか減らない．すなわち，$w(T)\geqq w(T')-1$ となる．同様に，v_{n+1}，v_{n+2} のうちどちらか一つは S' に入っているので，$|S|\leqq|S'|-1$ が成り立つ．v_{n+1} と v_{n+2} を取り去ることにより枝数は減少するので，$\sum_{Q_i}\left\lfloor\frac{edge(Q_i,T)}{2}\right\rfloor\leqq\sum_{Q'_i}\left\lfloor\frac{edge(Q'_i,T')}{2}\right\rfloor$ となる．従って，$w(T)\geqq w(T')-1>|S'|-1+\sum_{Q'_i}\left\lfloor\frac{edge(Q'_i,T')}{2}\right\rfloor\geqq|S|+\sum_{Q_i}\left\lfloor\frac{edge(Q_i,T)}{2}\right\rfloor$ となり，分割 Π は (2) を満たすことがわかる．

京都大学 KYOTO UNIVERSITY
A Self-archived copy in
Kyoto University Research Information Repository
https://repository.kulib.kyoto-u.ac.jp
京都大学学術情報リポジトリ KURENAI 紅
Kyoto University Research Information Repository

次に，Π' の元で $v_n \in S'$ の場合を考える．v_n と v_{n+2} は同じ頂点につながっているので，*NS2F*への属性を考えたとき，全く同じ役割をしていることがわかる．すなわち，v_n と v_{n+2} は対称である．従って，この場合は，Π' による V' の分割で，v_n と v_{n+2} の属性を入れ替えても式 (1) を満たす．よって，$v_{n+2} \in S'$ の場合と同様に考えることができ，分割 Π を構成することができる．

場合（B）：この場合は，更に細かく三つの場合分けをして考える．(B-1) v_{n+1} と v_{n+2} が R' に入っている場合．(B-2) v_n と v_{n+1} が R' に入っている場合．(B-3) v_n と v_{n+2} が R' に入っている場合．

場合（B-1）：この場合は，Π' と全く同様に Π を構成すればよい．$w(T) = w(T')$，$|S| = |S'|$，$\sum_{Q_i} \left\lfloor \frac{edge(Q_i,T)}{2} \right\rfloor \leqq \sum_{Q'_i} \left\lfloor \frac{edge(Q'_i,T')}{2} \right\rfloor$ となり，分割 Π が式 (2) を満たしていることが容易にわかる．

場合（B-2）：v_n と v_{n+2} の対称性を用いて（B-1）の場合と同様に考えることができる．

場合（B-3）：$v_{n+1} \in R'$ の場合は，(B-1) または (B-2) の場合に含まれているので，ここでは，Π' のもとで $v_n \in R'$，$v_{n+1} \in T'$，$v_{n+2} \in R'$ となっている場合のみを考えればよい．まず，v_{n+2} につながっている頂点のうち，T' に割り当てられているものが少なくとも一つある場合を考える．それを v_j とする．この場合は，Π を Π' と全く同じにする．すると，$|S| = |S'|$，$w(T) = w(T') - 1$ となっていることは容易にわかる．また，$v_{n+2}(\in R')$ と T' の頂点を結ぶ枝は少なくとも二つ（(v_{n+2}, v_{n+1}) と (v_{n+2}, v_j)）あるので，$\sum_{Q_i} \left\lfloor \frac{edge(Q_i,T)}{2} \right\rfloor \leqq \sum_{Q'_i} \left\lfloor \frac{edge(Q'_i,T')}{2} \right\rfloor - 1$ となっている．従って，Π は (2) を満たしている．よって，今後は，v_{n+2} とつながっている頂点はどれも T' に入っていない場合を考える．v_{n+2} につながっている頂点のうち，R' に割り当てられているものが少なくとも一つあるとする．この場合も Π を Π' と同じにする．v_n，v_{n+2} はともに R' の頂点で，ともに他の R' の頂点とつながっているので，v_n と v_{n+2} は同じ連結成分に入っている．二つの枝 (v_n, v_{n+1}) と (v_{n+1}, v_{n+2}) は $\delta(G)$ にはあるが，G にはないので，$\sum_{Q_i} \left\lfloor \frac{edge(Q_i,T)}{2} \right\rfloor \leqq \sum_{Q'_i} \left\lfloor \frac{edge(Q'_i,T')}{2} \right\rfloor - 1$ である．また，$w(T) = w(T') - 1$ であるため，Π が (2) を満たしていることがわかる．v_{n+2} につながっている頂点が v_{n+1} 以外はすべて S' の頂点なら，v_n 以外は Π' と同じ分割にして，v_n を T に割り当てれば，式 (2) を満たす．

場合（C）：（A），（B）で取り扱われなかったのは，以下の四つの場合のみである．(C-1) $v_n \in T'$，$v_{n+1} \in T'$，$v_{n+2} \in R'$．(C-2) $v_n \in T'$，$v_{n+1} \in R'$，$v_{n+2} \in T'$．(C-3) $v_n \in R'$，$v_{n+1} \in T'$，$v_{n+2} \in T'$．(C-4) $v_n \in T'$，$v_{n+1} \in T'$，$v_{n+2} \in T'$．いずれの場合も，Π を Π' と同じにすれば (2) が満たされていることを，これまでと同様にしてチェックすることができる．□

［系 1］ NP$\neq$coNP ならば，*NH*の NP 完全近似 *NS2F*は最適ではない．

5. δ に対する条件の緩和

定理 1 において，$\overline{B}$ が δ で閉じているという条件が必要であった．しかし，この条件は強すぎるように思える．δ は長さ増加関数である．従って，$|g| = n$ なら $|\delta^{2^n}(g)| > 2^n$ である．$f = \delta^{2^n}(g)$ とする．A と $\overline{A}$ は δ で閉じているので，g が A に入るかどうかをチェックすることにより，f が A に入るかどうかを判定できる．この場合，f のサイズが十分に大きいので，$g \in A$ の判定に指数時間かかっても $f \in A$ の判定は多項式時間でできることになる．我々は，近似 B をできるだけ大きくしようとしているので，B がこのような f を含むことは十分に考えられる．従って，$g \in A - B$ に対して，十分大きな i をとっても $\delta^i(g) \notin B$ であるという条件は自然ではない．本章では，この条件を，より自然な条件に弱めても，定理 1 が成り立つことを示す．

S を集合，$p(n)$ を実数から実数への関数とする．列 f に対して，ある列 g と整数 $i \geqq p(|g|)$ が存在し，$f = \delta^i(g)$ と書けるとき，f は関数 $p(n)$ に関して成熟であると言う．任意の $f \in S$ に対して，f が $p(n)$ に関して成熟でないならば $\delta(f) \in S$ であるとき，集合 S は関数 $p(n)$ に関して関数 δ で閉じていると言う．f が $p(n)$ に関して成熟ならば，$\delta(f)$ も $p(n)$ に関して成熟であることに注意されたい．

［定理 3］ 定理 1 において，$\overline{B}$ が δ で閉じているという条件を，$\overline{B}$ が，ある多項式 $p(n)$ に関して δ で閉じているとしても定理は成り立つ．

（証明） 本定理の証明は，定理 1 の証明とほぼ同様に行うことができる．但し，定理 1 の証明中に使われた多項式時間変換 α' と集合 D，E を以下のように変更する．

（i） $E = \{f \mid$ (a) $\delta^{-1}(f) \in B$ かつ f は成熟でな

京都大学 KYOTO UNIVERSITY
A Self-archived copy in
Kyoto University Research Information Repository
https://repository.kulib.kyoto-u.ac.jp
京都大学学術情報リポジトリ KURENAI 紅
Kyoto University Research Information Repository

いまたは (b) $\delta^{-1}(f) = \Lambda$ かつ $\psi(f) \in B\}$.

(ii) $D = \{\delta(f) \mid f \in C$ かつ f は成熟でない$\} \cup \{f \mid f \in C$ かつ f は成熟である$\}$.

(iii) $\alpha(x)$ が成熟ならば，$\alpha'(x) = \delta \cdot \alpha(x)$，$\alpha(x)$ が成熟でないならば，$\alpha'(x) = \alpha(x)$.

この場合も，定理1と同様，$(B \cup E) \cap D = \emptyset$ となることに注意されたい.

今回は，定理1の場合と違い $C - D$ が有限集合である可能性がある．この場合，定理1で使った，$C - D$ の要素に δ^{-1} を繰り返し適用させていく方法では，$E - B$ が無限であることを示すことはできない．しかし，$C - D$ が有限であるならば，以下の理由により $A - (B \cup C \cup D)$ が無限集合であることがわかる．$A - (B \cup C \cup D)$ が有限であると仮定する．集合 C，D の定義から，$|D - C| \leqq |C - D|$ であることがわかる．$C - D$ が有限集合なので，$D - C$ も有限集合である．従って $A - (B \cup (C \cap D))$ は有限集合である．また，$C_0 \cap D = \emptyset$ であるため，$C_0 - B$ は有限集合であり，$B \in \mathrm{NP}$ なので，$B \cup C_0 \in \mathrm{NP}$ である．集合 B' を $B' = \{\delta^i(f) \mid f \in B \cup C_0, i \geqq 0\}$ と定義する．列 f が与えられたとき，$B \cup C_0$ の要素と，δ の適用回数をゲスすることにより $f \in B'$ を多項式時間で判定できる．従って $B' \in \mathrm{NP}$ である．B' の定義より $B' \supseteq B \cup (C \cap D)$ であり，$A - (B \cup (C \cap D))$ は有限なので，$A - B'$ は有限集合である．$B' \in \mathrm{NP}$ なので $A \in \mathrm{NP}$ であり，NP=coNP となり仮定に矛盾する．従って $A - (B \cup C \cup D)$ は無限集合である．よって，$C - D$ が有限集合である場合，命題1，2と同様にして，この無限集合 $A - (B \cup C \cup D)$ の要素に ψ を適用させることにより $E - B$ が無限集合であるという結果を得ることができる． □

6. むすび

本論文では，coNP 集合を NP 完全集合で近似する際の，最適近似の存在性について議論を行った．その結果，比較的弱い条件の元で，最適近似が存在しないことがわかった．また，そのような条件を満たす coNP 完全集合と，その NP 完全近似の例を示した．

謝辞 本研究は，文部省科学研究費補助金によるものである．

文 献

[1] Y. Asahiro, K. Iwama, and E. Miyano, "Random generation of test instances with controlled attributes," Proc. Second DIMACS Challenge Workshop, eds. M. Trick and D. Johnson, American Mathematical Society, 1995.

[2] J. Balcazar, J. Diaz, and J. Gabarro, "Structural Complexity II," Springer-Verlag, Berlin, 1989.

[3] D. Bauer, S.L. Hakimi, and E. Schmeichel, "Recognizing tough graphs is NP-hard," Discrete Applied Mathematics, vol.28, pp.191–195, 1990.

[4] L. Berman and J. Hartmanis, "On isomorphism and density of NP and other complete sets," SIAM J. Comput., vol.6, pp.305–322, 1977.

[5] V. Chvátal, "Hamiltonian cycles," The traveling salesman problem, John Wiley & Sons Ltd., pp.403–429, 1985.

[6] A. Haken, "The intractability of resolution," Theoretical Computer Science, vol.39, pp.297–308, 1985.

[7] K. Iwama and E. Miyano, "Security of test-case generation with known answers," Proc. AAAI Spring Symposium Series, pp.85–91, 1993.

[8] K. Iwama and E. Miyano, "Intractability of read-once resolution," Proc. 10th IEEE Conference on Structure in Complexity Theory, pp.29–36, 1995.

[9] K. Iwama and E. Miyano, "Better approximations of non-Hamiltonian graphs," Discrete Applied Mathematics, vol.81, pp.239–261, 1998.

[10] P. Orponen, D. Russo, and U. Schöning, "Optimal approximations and polynomially levelable sets," SIAM J. Comput., vol.15, pp.399–408, 1986.

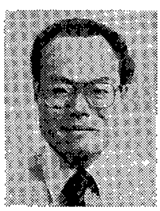
[11] T. Pitassi and A. Urquhart, "The complexity of Hajós calculus," Proc. 33rd IEEE Symp. on Foundations of Computer Science, pp.187–196, 1992.

[12] L. Sanchis, "Generating hard and diverse test sets for NP-hard graph problems," Discrete Applied Mathematics, vol.58, no.1, pp.35–66, 1995.

(平成9年8月28日受付，12月25日再受付)

宮崎 修一

平5九州大・工・情報卒．平7同情報工学専攻了．平10同博士後期課程了．工博．現京大・工・情報助手．計算の複雑さ理論の研究に従事．



岩間 一雄 (正員)

昭48京大・工・電気卒．昭53同大大学院博士課程了．工博．昭54京産大・理・計算機科学科・講師．昭57同助教授．昭58～59カリフォルニア大学バークレー客員準教授．平2九大・工・情報・助教授，平4同教授を経て，平9京大・工・情報・教授．計算の複雑さの理論，並列アルゴリズム等の研究に従事．